

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Jurisprudencia

**Inteligencia artificial y la institución motivacional: viabilidad de
la implementación de sistemas para decisiones judiciales**

Matías Nicolás Castro Benítez

Jurisprudencia

**Trabajo de fin de carrera presentado como requisito para la
obtención del título de Abogado**

Quito, 17 de abril de 2025

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Nombres y apellidos: Matías Nicolás Castro Benítez

Código: 00324465

Cédula de identidad: 1724060999

Lugar y Fecha: Quito, 17 de abril de 2025

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETheses>.

UNPUBLISHED DOCUMENT

Note: The following capstone Project is available through Universidad San Francisco de Quito USFQ institutional repository. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETheses>.

*Agradezco a mis padres Paola y Germán por haber confiado siempre
en mí y haberme brindado el más dulce de los apoyos.*

*A mis amigos, Daniel y Duván por acompañarme y resguardarme
estos días con su cálida amistad.*

*A mi tutor Juanse, que fue un extraordinario guía y que me brindó más
ayuda de la que podría haber pedido.*

*A Juandi por su invaluable ayuda en la formación de este trabajo y
por su vastísimo aporte intelectual.*

INTELIGENCIA ARTIFICIAL Y LA INSTITUCIÓN MOTIVACIONAL: VIABILIDAD DE LA IMPLEMENTACIÓN DE SISTEMAS PARA DECISIONES JUDICIALES¹

ARTIFICIAL INTELLIGENCE AND JUDICIAL MOTIVATION: THE FEASIBILITY OF IMPLEMENTING SYSTEMS FOR JUDICIAL DECISIONS

Matías Nicolás Castro Benítez²
nicolascastro20052002@gmail.com

RESUMEN

El avance de la Inteligencia Artificial, particularmente los Grandes Modelos de Lenguaje, plantea su aplicación en decisiones judiciales. Estas deben cumplir con la motivación, garantía constitucional que exige justificación racional, transparencia y consideración de valores, trascendiendo la mera coherencia formal. Este trabajo analiza la viabilidad de que los GML generen autónomamente decisiones motivadas. Se contrastan los requisitos de la motivación (lógica, axiológica, funcional) con el funcionamiento y limitaciones de los GML (procesamiento estadístico-sintáctico, opacidad). Se concluye que, por su incapacidad de comprensión semántica y razonamiento axiológico, y el problema de la "caja negra" que impide la transparencia y el control, los GML actuales no pueden satisfacer las finalidades sustantivas de la motivación. Aunque generen textos superficialmente plausibles, sus barreras cognitivas y epistémicas los hacen inviables para esta función judicial autónoma en el Estado de Derecho.

PALABRAS CLAVE

Motivación Judicial, Inteligencia Artificial, Grandes Modelos de Lenguaje, Transparencia, Justificación Judicial

ABSTRACT

The advancement of Artificial Intelligence (AI), particularly Large Language Models (LLMs), raises its application in judicial decisions. These must comply with motivation, a constitutional guarantee demanding rational justification, transparency, and consideration of values, transcending mere formal coherence. This work analyzes the viability of LLMs autonomously generating motivated decisions. The requirements of motivation (logical, axiological, functional) are contrasted with the functioning and limitations of LLMs (statistical-syntactic processing, opacity). It is concluded that, due to their inability for semantic understanding and axiological reasoning, and the "black box" problem hindering transparency and control, current LLMs cannot satisfy the substantive purposes of motivation. Although they may generate superficially plausible texts, their cognitive and epistemic barriers make them unviable for this autonomous judicial function in the Rule of Law.

KEYWORDS

Judicial Motivation, Artificial Intelligence, Large Language Models, Transparency, Judicial Justification

¹ Trabajo de titulación presentado como requisito para la obtención del título de Abogado. Colegio de Jurisprudencia de la Universidad San Francisco de Quito. Dirigido por Juan Sebastián Baquero Correa.

² © DERECHOS DE AUTOR: Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas. Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Fecha de lectura: 17 de abril de 2025

Fecha de publicación: 17 de abril de 2025

SUMARIO

1. Introducción.- 2. Marco teórico.- 3. Estado del Arte.- 4. Marco Normativo.- 5. La figura motivacional: Origen, estructura y función.- 5.1. Necesidad e importancia motivacional en las decisiones judiciales.- 5.2. Motivación como garantía de transparencia y legitimidad.- 5.3. Las teorías justificativas.- 5.4. Motivación más allá de esquemas justificativos.- 6. Inteligencia Artificial.- 6.1 Inteligencia Artificial y Grandes Modelos de Lenguaje.- 6.2 Proceso interno de razonamiento de los Grandes Modelos de Lenguaje.- 6.3 Transparencia y el problema de caja negra.- 6.4. Inteligencia Artificial General como una solución ideal.- 7. Capacidad motivacional de Grandes modelos de Lenguaje.- 7.1 Entendimiento Artificial vs Entendimiento Humano.- 7.2 ¿Puede un Gran Modelo de Lenguaje motivar?.- 7.3 Limitaciones axiológicas del Gran Modelo de lenguaje.- 7.4 Comparativa entre decisiones tomadas por un juez y decisiones tomadas por un algoritmo.- 8 Conclusiones

1. Introducción

En 1950, Turing publicó su ensayo "*Computing Machinery and Intelligence*", el cual comenzaba con una interrogante cuya importancia se haría sentir en el futuro: “¿Pueden las máquinas pensar?”³. Esta pregunta marcó un hito en la matemática y la ciencia computacional que a su vez impulsó una carrera de innovación tecnológica de la cual nacerían los modelos modernos de “Inteligencia Artificial”. Estos nuevos modelos informáticos y su facilidad de acceso implican un cambio de paradigma para la práctica legal.

³ Turing, Alan, “Computing Machinery and Intelligence”, *Mind* 49, (1950): 433-460.

En el ámbito jurídico estas herramientas de procesamiento de información han contribuido a la optimización y automatización de tareas de gran volumen, al punto que permiten el análisis y clasificación de grandes cantidades de información. Un ejemplo de ello es el sistema PretorIA empleado por la Corte Constitucional Colombiana, mismo sistema que examina y ayuda en el proceso de selección de acciones de tutela⁴, permitiendo un dinamismo en la administración de justicia sin precedentes.

Ahora bien, la problemática surge cuando trasladamos el uso de estos sistemas hacia la potestad de dirimir controversias que posee un juez, puesto que surgen una serie de dudas respecto a las decisiones judiciales impartidas haciendo uso de los sistemas de Inteligencia Artificial. Particularmente, si este tipo de decisiones mantienen la integridad motivacional de la que deben estar revestidas.

La propia naturaleza de estos sistemas informáticos implica varios problemas impregnados en su diseño que desembocan en dilemas de ética jurídica y filosofía del Derecho que merecen ser abordados. En particular, su transparencia, existencia de sesgos y aquellos derivados de la propia naturaleza de la argumentación y lógica jurídica. En tal sentido, el presente trabajo pretende examinar hasta qué punto una decisión judicial que haya sido auxiliada por un sistema de Inteligencia Artificial goza de plena motivación y justificación conforme al Derecho. Sin embargo, su enfoque no se limita a un análisis de si estos sistemas podrían cumplir un *test* de suficiencia motivacional, sino también de la capacidad de estos sistemas de justificar las decisiones, razonar jurídicamente y cumplir las finalidades legitimadoras de la motivación en un Estado Constitucional.

Para efectos de la presente exposición nos centraremos en el uso de un modelo en específico de Inteligencia Artificial: los Grandes Modelos de Lenguaje o *Large Language Model* (LLM). Estos son sistemas que replican el lenguaje humano de manera predictiva y poseen una arquitectura de *Generative Pretrained Transformer* o Transformador Generativo Preentrenado (GPT). Este programa se encarga de recibir instrucciones y generar una respuesta conforme a sus parámetros internos y entrenamiento, mismo que es particularmente

⁴ Juan Corvalán, “Inteligencia Artificial GPT-3, PretorIA y oráculos algorítmicos en el Derecho,” *International Journal of Digital Law*, no. 1 (2020), 43-46.

importante para el campo del Derecho dado que su capacidad de procesamiento del lenguaje humano lo vuelve ideal para el desarrollo jurídico.

La pregunta, por tanto, surge sobre la capacidad de la máquina para tomar decisiones que cumplan con las necesidades constitucionales que son exigidas para que formalmente exista una motivación, y si esas decisiones por sí mismas pueden gozar de plenitud motivacional más allá de lo aparente a primera lectura. Por ello la metodología planteada en el presente trabajo es la metodología exegética cualitativa, es decir, se centrará en la revisión de la doctrina jurídica y técnica para determinar el alcance de la figura motivacional en relación con los modelos de Inteligencia Artificial.

2. Estado del Arte

Esta sección revisa las posturas doctrinales más relevantes sobre la motivación judicial y analiza los debates emergentes que la nueva era tecnológica plantea a este y otros conceptos jurídicos fundamentales.

Dentro de la doctrina fundamental sobre motivación, Calamandrei subraya su rol como contrapeso de la arbitrariedad judicial, pues el deber de motivar transforma la decisión judicial donde deja de ser un potencial acto de voluntad autoritaria para convertirse en una manifestación de racionalidad, donde se exponen las razones que la sustentan y la distinguen del simple poder discrecional.⁵

En este sentido, Ricaurte sostiene que los fallos emitidos por la autoridad jurisdiccional deben contar con una motivación suficiente, es decir, seguir un estándar que además de garantizar la relación lógica entre supuestos fácticos y normativos, debe también alejarse de un examen de corrección objetivo. De este modo a la luz del Derecho una decisión motivada gozará de indudable robustez en su argumentación, legitimando la resolución.⁶ Ahora bien, una vez discutidas algunas posturas relevantes que introducen la idea de la motivación -y que serán exploradas con mayor detalle en secciones subsiguiente-

⁵ Piero Calamandrei, *La motivación de la sentencia Civil*, (Madrid: Ediciones Jurídicas Europa-América, 1960), 115.

⁶ Catherine Ricaurte, “Derecho a la motivación: Análisis a partir de la sentencia 1158-17-EP/21 de la Corte Constitucional del Ecuador”, *Revista Cálamo* no. 18 (2023), 31–44.

corresponde mencionar algunas de las posturas que exploran las implicaciones de la innovación tecnológica con el Derecho.

Cofone sostiene que la facultad de juzgar que posee un juez es, de momento, exclusiva de los seres humanos pero el uso de algoritmos generativos puede ser útil para el juzgador al momento de emitir una decisión. No obstante, reconoce las dificultades de transparencia en cuanto a la lógica y el razonamiento empleado por los algoritmos generativos, creando un escenario en donde las partes implicadas no conocen la manera en la que una decisión fue tomada, impidiendo así un claro proceso resolutorio.⁷

Asimismo, Crotoft considera que el rol de algoritmos predictivos en la justicia puede contribuir de gran manera en la optimización y eficiencia al momento de impartir justicia por el magistrado, pero también destaca los riesgos del modelo híbrido entre juez y máquina, pues contempla situaciones en que el algoritmo en lugar de servir como una ayuda para el juez termina amplificando las deficiencias subyacentes en su juicio, influyendo en la formación de vicios motivacionales dentro de las resoluciones judiciales asistidas por sistemas.⁸

Por otro lado, Montañez argumenta que los modelos de Inteligencia Artificial permiten al magistrado tomar decisiones que sean mayormente basadas conforme al derecho y dejando de lado las subjetividades inherentes al juzgador en su calidad de ser humano, apostando por una mirada objetiva hacia la nueva administración de justicia, sin desplazar completamente al juez como el órgano jurisdiccional, puesto que seguirá siendo una figura crucial en la supervisión de la máquina, adoptando una postura en que la máquina se ve como una extensión o evolución en el proceso motivacional para alcanzar un ideal de objetividad argumentativa.⁹

Frente a este tipo de discrepancias sobre los beneficios y dificultades del uso de Inteligencia Artificial en la administración de justicia, la presente investigación pretende

⁷ Ignacio Cofone, “AI and Judicial Decision-Making”, en *Artificial Intelligence and the Law in Canada*, ed. por Florian Martin-Bariteau & Teresa Scassa. (Toronto: LexisNexis Canada, 2021), capítulo 13.

⁸ Rebecca Crotoft, “‘Cyborg Justice’ and the risk of Technological–Legal Lock-In”, *Columbia Law Review Forum* 119, (2019): 233-251.

⁹ Cristian Montañez, “Subjetividad de las máquinas en toma de decisiones judiciales”, *Foro: Revista De Derecho* 36,(2021), 9–27.

abonar a la cuestión a partir del análisis de lo que exige constitucionalmente la motivación y las limitaciones de la Inteligencia Artificial para razonar.

3. Marco teórico

La discusión doctrinaria concerniente al concepto de motivación ha demostrado ser extensa y gozar de distintos postulados que pretenden encasillar esta exigencia judicial en una definición concreta. De este modo, surgen teorías clásicas como la teoría psicologista de la motivación, la cual determina que esta es la expresión lingüística de los motivos o factores causales que explican la convicción del juez al tomar una decisión.¹⁰ Al mismo tiempo, existen otras teorías sobre la motivación como la concepción racionalista, la cual enmarca a las decisiones motivadas como un proceso justificativo de razones que culminan en una decisión.¹¹

Ambas posturas basan su tesis en que la motivación es parte de la expresión lingüística del administrador de justicia encaminada a exteriorizar su proceso intelectual. Sin embargo, estas se diferencian en que mientras la concepción psicologista busca explicar los motivos del juzgador, la concepción racionalista pretende que los jueces brinden razones suficientes para justificar las decisiones.

Si bien estas posturas buscan solventar el problema conceptual de la motivación, varios sectores de la doctrina jurídica han criticado la enmarcación motivacional como una mera exteriorización realizada por el agente judicial, puesto que resulta una visión reduccionista que limita el alcance de la figura de la motivación a un formalismo duro.¹²

Ahora bien, la doctrina adoptada por Calamandrei sostiene que la motivación comprende dos facetas: Lógica y axiológica. La esfera lógica bajo la que se encuentra envuelta la motivación se describe La exposición de las bases del silogismo legal que conducen directamente a la parte dispositiva del fallo judicial.¹³ Por otro lado, la faceta axiológica de la motivación atiende a cuestiones subyacentes al ejercicio lógico realizado por el juzgador,

¹⁰ Jordi Beltrán, “Apuntes sobre el Concepto de Motivación de las Decisiones Judiciales”, *Isonomía* 34 (2011), 89.

¹¹ Colomer Hernández, *La motivación de las sentencias: sus exigencias constitucionales y legales*, (Valencia: Tirant lo Blanch, 2003), 31

¹² Juan Igartura Salaverría, *El razonamiento en las decisiones judiciales*, (Lima-Bogotá, Palestra, 2009), 30

¹³ Piero Calamandrei, *La motivación de la sentencia Civil*, 116.

y en su lugar pretende precautelar y consagrar los valores constitucionales del Estado de Derecho. Dicho de otro modo, la postura del autor presenta la motivación judicial de una forma integral, que no se limita a lo meramente formal. Enfatiza su importancia sustantiva, sosteniendo que, al basarse en la justicia y principios del Derecho, las decisiones de los jueces adquieren mayor fuerza y legitimidad.

Una vez exhibida la base teórica que atañe a la motivación, es pertinente la revisión de las posturas doctrinales relevantes tanto de la figura motivacional, como de la aplicación de los sistemas al proceso resolutorio del juzgador.

4. Marco normativo

En esta sección se enunciará el marco legal tanto nacional como internacional que construye tanto el marco motivacional como el régimen de control sobre los sistemas de Inteligencia Artificial. Dado que la naturaleza del presente trabajo se enfoca en el plano teórico y en algunos de los posibles problemas de la Inteligencia Artificial para motivar y justificar las decisiones, la norma no será el punto central de la investigación. Sin embargo, es pertinente destacar la normativa de mayor relevancia para el desarrollo de la discusión.

En lo que respecta a la normativa local relevante, es pertinente mencionar el derecho al debido proceso, particularmente el derecho a la defensa en su garantía de motivación reconocida en el artículo 76 numeral .7 literal .1) de la Constitución de la República del Ecuador, que dispone lo siguiente:

Las resoluciones de los poderes públicos deberán ser motivadas. No habrá motivación si en la resolución no se enuncian las normas o principios jurídicos en que se funda y no se explica la pertinencia de su aplicación a los antecedentes de hecho. Los actos administrativos, resoluciones o fallos que no se encuentren debidamente motivados se considerarán nulos. Las servidoras o servidores responsables serán sancionados.¹⁴

De este modo, esta norma establece la obligación constitucional de que las resoluciones emitidas por el poder público se encuentren motivadas y, en particular, garantiza el derecho a la defensa, ya que las partes procesales no podrían entablar alegaciones contra las decisiones, ni impugnarlas de manera efectiva si no conocen sus razones.

¹⁴ Constitución de la República del Ecuador, R.O. 449, 20 de octubre de 2008, reformada por última vez el 30 de mayo de 2024.

Bajo este mismo principio el artículo 130 del Código Orgánico de la Función Judicial¹⁵ establece las facultades de los jueces, donde su cuarto numeral detalla la obligación de motivar las resoluciones emitidas por autoridades jurisdiccionales. La motivación por su parte se encuentra igualmente definida en el artículo 89 del Código Orgánico General de Procesos,¹⁶ mismo que establece que las sentencias gozarán de motivación cuando sus decisiones se encuentren fundadas con razonamientos fácticos y jurídicos que estén encaminados a la apreciación de la prueba y aplicación del derecho.

Respecto a la jurisprudencia constitucional ecuatoriana, es relevante destacar la sentencia No. 1158-17-EP¹⁷, en donde la Corte Constitucional del Ecuador analiza la definición y el alcance de la motivación judicial. En esta sentencia constitucional, la Corte define un cambio del marco jurídico alrededor de la garantía motivacional. Así es como se aparta del test de motivación y se centra en la suficiencia motivacional, aclarando que esta garantía no pretende la corrección de las decisiones judiciales.

En lo que respecta a la normativa aplicable a la temática de la inteligencia artificial es pertinente reconocer que el esquema normativo sigue en desarrollo. Localmente en Ecuador no se encuentra reglado el uso de la Inteligencia Artificial dentro de las potestades del poder judicial. Sin embargo, destaco particularmente el reglamento promulgado por la Unión Europea y que entró en vigor en el año 2024. Este cuerpo legal establece un esquema normativo que regula y estandariza el ecosistema pertinente a la Inteligencia Artificial, sus dificultades y posibles aplicaciones desde una perspectiva jurídica. Este Reglamento es, la norma más relevante en la actualidad cuando se explora la estandarización jurídica de la inteligencia artificial y ofrece tanto definiciones como medidas de regulación innovativas que serán tomadas en consideración para el presente trabajo.¹⁸

¹⁵ Artículo 130, Código Orgánico de la Función Judicial [COFJ]. R.O. 544, 09 de marzo de 2009, Reformado por última vez 29 de marzo de 2023.

¹⁶ Artículo 89, Código Orgánico General de Procesos [COGEP]. R.O. 506, 22 de mayo de 2015, Reformado por última vez 05 de enero de 2024

¹⁷ Causa No. 1158-17-EP, Corte Constitucional, 20 octubre de 2021.

¹⁸ Reglamento 2024/1689, Parlamento Europeo, Consejo [por el que se establecen normas armonizadas en materia de inteligencia artificial], publicado en el Diario Oficial de la Unión Europea el 12 de julio de 2024.

5. La figura motivacional: Origen, estructura y función

Una parte fundamental del enfoque adoptado por el presente ensayo es examinar el alcance, la naturaleza y la importancia del fenómeno de la motivación judicial, a fin de comprender su relevancia dentro del Estado de Derecho. Estas ideas nos permitirán en las secciones subsiguientes realizar un análisis de la Inteligencia Artificial y sus capacidades para adentrarse a la esfera motivacional.

Para Calamandrei, “La motivación constituye el signo más importante y típico de la ‘racionalización’ de la función jurisdiccional”¹⁹. De esta manera para el autor italiano la motivación incluida en los fallos realizados por administradores de justicia es un ejercicio realizado con el fin de evitar que la sentencia sea un mandato autoritario ordenado por el juez de forma arbitraria, es decir, la exigencia de motivar una decisión busca dotar a la sentencia de una justificación que plasme el proceso de razonamiento lógico efectuado por el juzgador, de modo que permita a los administrados de justicia verificar que la resolución adoptada es justa.

5.1. Necesidad e importancia motivacional en las decisiones judiciales

Realizando una revisión histórica a la manera en que las decisiones judiciales eran emitidas, si nos remontamos al derecho romano, originalmente la motivación de las sentencias no era una exigencia pues los administradores de justicia eran los conocedores de los textos legales y, dado su estatus como nobles eran los únicos capaz de ejercer la capacidad de juzgar o *iuris dictio*.²⁰ No obstante, con la redacción de las 7 partidas introducidas a lo largo del siglo XIII en el reinado de Alfonso X, se consagró en su tercera partida la obligación de identificar la causa que llevó al administrador de justicia a tomar una decisión.²¹ Siendo este el primer antecedente histórico de un deber motivacional.

Sin embargo, la figura de la motivación moderna se remonta a la Revolución Francesa de finales del siglo XVIII, en donde se plantea un cambio estructural en los sistemas de justicia, surgiendo así el Código de Procedimiento Civil de 1807. En este cuerpo normativo,

¹⁹ Piero Calamandrei, *La motivación de la sentencia Civil*, 115

²⁰ Osvaldo Gozaíni, *Tratado de Derecho Procesal Constitucional Latinoamericano. Tomo I*, (Buenos Aires: La Ley, 2014), 703

²¹ Id, 704

nace la motivación como una herramienta que permitía la revisión de decisiones impugnadas por parte de los jueces superiores. Este acontecimiento marcó un precedente que sería adoptado por otras naciones europeas en los años venideros y que, consecuentemente, refinarían la institución motivacional²².

Sin embargo, no fue sino hasta años posteriores a la Segunda Guerra Mundial con la formación de nuevas constituciones democráticas que el deber de motivar pasaría a considerarse un principio constitucional estrechamente vinculada al debido proceso, fortaleciendo la formación de los Estados con democracias constitucionales. Esto es así dado que la motivación es vista como una garantía del Estado de Derecho que evita los modelos de Estados autoritarios y la administración arbitraria de justicia. Así es como empieza la concepción que vincula la motivación y el poder.²³

El antiguo sistema español ilustra bien de qué quería distanciarse el modelo francés: en España, estaba prohibido que los jueces motivaran sus fallos, ya que justificar una decisión judicial se consideraba un acto que demostraba la autoridad del Rey, quien era la fuente de la justicia, guiándose bajo la máxima *Rex non potest peccare*.²⁴

La relación entre la motivación y el proceso democrático es el punto culmen del desarrollo legislativo histórico brevemente descrito en esta sección, donde la exigencia constitucional de motivar las sentencias de los magistrados tiene el objetivo de transparentar y plasmar de manera comprensible su ejercicio del poder conferido. Es decir, la motivación genera una suerte de proximidad con la ciudadanía, esta proximidad permite conocer plenamente el proceso justificativo adoptado por el juzgador para su resolución²⁵ y de este modo el control al que se somete goza de mayor amplitud a comparación del que podría ejercer otro juez, siendo este un tipo de control externo sometido a la opinión pública.

Se trata pues, de la antítesis de la mencionada máxima *Rex non peccare*, donde se despoja el poder de carácter incuestionable en los sistemas autoritarios para revestir las decisiones de legitimidad ante la sociedad. Es de este modo que la necesidad de contar con

²² Michele Taruffo, *Proceso y Decisión Lecciones mexicanas de Derecho Procesal*, (Madrid: Marcial Pons, 2012), 87-90

²³ Michele Taruffo, *Proceso y Decisión Lecciones mexicanas de Derecho Procesal*, 94.

²⁴ Ignacio Colomer Hernández, *La motivación de las sentencias: sus exigencias constitucionales y legales*, 60

²⁵ Jordi Ferrer Beltrán, “Apuntes sobre el concepto de motivación de las decisiones judiciales”, 96-102.

decisiones motivadas se encuentra justificada en virtud de preservar la integridad del Estado de Derecho y del proceso democrático.

5.2. Motivación como garantía de transparencia y legitimidad

La revisión histórica efectuada deja en evidencia que gran parte de la importancia que adquiere la garantía motivacional es a través de su función como un medio que permite a la sociedad conocer la manera en que se ejerce el poder jurisdiccional. Para Couture, la motivación requiere que el juez estructure su actividad intelectual respecto a la causa, garantizando que el fallo sea un acto reflexivo fundado en el estudio particular de los hechos, y no una mera expresión de su voluntad discrecional.²⁶

Doctrinariamente se identifica una bipartición en las funcionalidades de la motivación: 1) Endoprocésal y 2) Extraprocésal. Esta división surge de la calidad que tengan las personas alrededor del ámbito jurisdiccional. Por el lado de la funcionalidad endoprocésal de la motivación, se centra en las partes que se encuentren en el interior del proceso y les permite a ellas conocer el razonamiento efectuado por el juzgador, de modo que puedan entender las razones que hayan llevado al magistrado a emitir una resolución a favor o en contra de sus intereses. Internamente, la motivación presenta a las partes una base lógica y argumentativa que posteriormente servirán para que los interesados dentro del proceso puedan allanarse a la resolución o ejercitar algún tipo de impugnación, a fin de precautelar su derecho a la defensa.

Por tanto, es menester identificar que la función endoprocésal de la motivación comprende dos niveles: (i) las partes y (ii) el poder judicial.²⁷ En palabras de Valenzuela, es una exigencia que posee un estrecho vínculo con las partes y el tribunal, pues el control que contempla se limita al mundo interno del proceso específico.²⁸

Por su parte, la función extraprocésal presenta un enfoque externo al proceso y las partes procesales. Se trata del sometimiento del ejercicio del poder judicial al escrutinio

²⁶ Eduardo Couture, *Fundamentos de Derecho Procesal Civil*, (Buenos Aires: B de F, 2007), 234.

²⁷ José Luis Castillo Alva, *Las funciones constitucionales del deber de motivar las decisiones judiciales*, 24.

²⁸ Gastón Fernando Valenzuela Pirotto, “Enfoque actual de la motivación de las sentencias. Su análisis como componente del debido proceso”, *Revista De Derecho* 21, (2020), 72-90.

social, este principio es referido por la Corte Interamericana de Derechos Humanos en el desarrollo del caso *Apitz Barbera y otros vs Venezuela* que realiza el siguiente análisis:

El deber de motivar las resoluciones es una garantía vinculada con la correcta administración de justicia, que protege el derecho de los ciudadanos a ser juzgados por las razones que el Derecho suministra, y otorga credibilidad de las decisiones jurídicas en el marco de una sociedad democrática.²⁹

Por ello, aunque el deber de motivación judicial posee una relación directa con el poder jurisdiccional y la tutela judicial efectiva, también es extendido hacia la generalidad de la ciudadanía, pues permite que los ciudadanos depositen su confianza en el sistema judicial, dotándolo así de plena legitimidad en el Estado democrático.

5.3 El proceso justificativo

Una vez identificada la importancia de la institución motivacional dentro del Estado de Derecho, es pertinente indagar en las características propias que llevan a que un juez pueda forjar una decisión que cuente con robustez motivacional. Para ello estudiar la manera en que un magistrado realiza su proceso justificativo antes de tomar una decisión.

La justificación en el ámbito judicial ha sido discutida por diversos autores que, en busca de establecer una definición, han planteado teorías que contemplan el proceso seguido por un órgano jurisdiccional para la emisión de un criterio. Comanducci sostiene que las teorías que interpretan la justificación jurídica encajan en tres enfoques: descriptivo, prescriptivo y teórico.³⁰ Estos enfoques representan diferentes perspectivas académicas sobre la justificación jurídica. El enfoque descriptivo busca entender y caracterizar la justificación como el proceso de razonamiento lógico interno que el juez emplea y que se manifiesta en la sentencia. Por su parte, el enfoque prescriptivo se orienta a establecer el estándar normativo o el nivel de calidad que debería alcanzar la justificación de un juez. Finalmente, el enfoque teórico propone modelos ideales de razonamiento que funcionan como un parámetro o referente para la práctica judicial del magistrado.³¹

²⁹ *Apitz Barbera y otros vs Venezuela*, Corte Interamericana de Derechos Humanos, 5 de agosto de 2008, párr. 77.

³⁰ Paolo Comanducci, *La interpretación jurídica*, 2011,

³¹ Ver: Jerzy Wroblewsky, *The Judicial Application of Law*, (Springer Science, 1992), 63

Ante esta diversidad de categorizaciones, Wroblewsky propone perspectivas que pretende englobar la generalidad de posturas y dividirlas en dos esquemas distintos: 1) Justificación interna y 2) Justificación externa.³² Estos enfoques son complementarios entre sí y el autor sostiene que son los niveles justificativos por la que toda decisión judicial debe atravesar para gozar de plenitud racional.³³

Partiendo por la justificación externa, esta aborda el razonamiento judicial desde una óptica ajena a un estudio centrado en el contenido y el silogismo judicial; Se centra en verificar la solidez del contenido de las premisas que apoyarán la conclusión adoptada.³⁴ Consiste, por tanto, en establecer y evaluar estos elementos fundamentales del caso. Su preocupación principal no es el 'ajuste lógico' entre premisas y conclusión (propio de la justificación interna), sino la corrección o validez ('rectitud') de esas premisas en sí mismas. En otros términos, busca garantizar la 'corrección material' de los fundamentos que luego empleará la justificación interna, asegurando así que el razonamiento no sea una mera secuencia lógica formal sin bases sólidas.³⁵ Lejos de ser ajena a la lógica, la justificación externa actúa como un complemento crucial que aporta mayor solidez y fundamento al proceso de razonamiento en su conjunto.

Por otro lado, la justificación interna es un método empleado por el juez para determinar que su decisión guarde coherencia con la norma y el cuerpo legal que conforma el ordenamiento jurídico³⁶, es una manera de control interno realizado para verificar que la resolución efectivamente sea conforme al Derecho. Por consiguiente, este modo justificativo es de tipo silogístico, en donde se ejercita el análisis lógico de las premisas para llegar a una determinada conclusión.³⁷ Es así como la justificación interna genera orden tanto en la coherencia lógica argumentativa, como en el seguimiento y cumplimiento de las reglas que comprenden el sistema jurídico, logrando un equilibrio entre ambos que legitima las decisiones.

³² Id, 62

³³ Id, 209

³⁴ Pedro Caballero, "Las decisiones judiciales: justificación y racionalidad", *Problema anuario de filosofía y teoría del derecho* 13 (2019), 80

³⁵ Juan Igartura Salaverría, *El razonamiento en las decisiones judiciales*, 30

³⁶ Jerzy Wróblewsky, *The Judicial Application of Law*, 63

³⁷ Manuel Atienza, *Curso de argumentación jurídica*, (Madrid, Trotta, 2013), 132.

5.4 Motivación más allá del proceso justificativo

Si bien la revisión de las teorías sobre la racionalización de las decisiones permite conocer la faceta lógica detrás del *iter* motivacional, la figura de la motivación no se limita a un examen de análisis lógico, Wróblewsky señala el carácter axiológico que se encuentra presente en las resoluciones judiciales que se encuentra desprendido de las reglas formales del racionalismo duro.³⁸ Este carácter alejado del ejercicio juicioso y metódico del razonamiento lógico responde a un sentido de practicidad argumentativa, en donde se incorporan los juicios valorativos y normativas de conducta en las decisiones judiciales, estas directrices a su vez conforman el marco general que seguirán los jueces al momento de motivar sus decisiones.³⁹ Demostrando así que la motivación no es estrictamente un problema racional esperando a ser resuelto, sino que atiene a cuestiones subyacentes más complejas.

Por consiguiente, la práctica motivacional de los jueces no está atado a una fórmula estática que mide la rectitud de las decisiones. Relacionado a este punto, Atienza encaja esta definición como la figura de motivación suficiente, donde le otorga valor a la razonabilidad de los argumentos esgrimidos.⁴⁰ Llevado al contexto local, la Corte Constitucional del Ecuador acoge la figura de la motivación suficiente en la sentencia No.1158-17-EP/21, donde argumenta que la institución motivacional no comprende un ajuste de argumentación correcta, sino que establece criterios mínimos que el proceso de justificación jurídica deberá seguir,⁴¹ siendo estos requisitos relevantes para la formulación y comunicación del *iter* mental del juzgador.

Aunado a esto, y como se expuso previamente en esta sección dedicada a la figura de la motivación, la institución motivacional y la obligación que poseen los jueces de emplearla para la emisión de sus fallos implica su importancia trascendental para dotar de legitimidad al poder judicial, siendo la herramienta por excelencia que permite a los ciudadanos, tanto dentro como fuera del proceso, conocer y determinar que la administración de justicia se lleva a cabo de manera legítima. Esta funcionalidad de transparencia no es una consecuencia

³⁸ Jerzy Wróblewsky, *The Judicial Application of Law*, 265.

³⁹ Manuel Atienza, *Curso de argumentación jurídica*, 133-134.

⁴⁰ Id, 136.

⁴¹ Causa No. 1158-17-EP/21, Corte Constitucional del Ecuador, 29 de octubre de 2021, párr. 26-28.

de la robustez racional que posea una resolución, sino que atiende a la importancia sustantiva que reviste a la decisión independientemente de su contenido formal.

De este modo es evidente que la motivación comprende un considerable número aristas y cumple un rol multifuncional dentro del estado de derecho, es por ello que es relevante trazar las características que debe poseer una sentencia motivada. Partiendo del hecho que la motivación tiene la finalidad de ofrecer una vista transparente desde la sociedad hacia el poder judicial, una decisión debe contar con el desarrollo adecuado que permita observar desde el exterior el camino argumentativo que siguió el magistrado. Además, debe comprender un desarrollo lógico que relacione los elementos tanto fácticos como normativos, atendiendo a la relación entre estos de una narrativa argumentativa clara, estructurada y coherente. En este sentido, una sentencia motivada no se limita a la exposición aislada de normas aplicables, sino que exige una integración interpretativa que justifique por qué determinadas normas resultan pertinentes y cómo se relacionan con los hechos probados en el caso.

Establecida la complejidad y la funcionalidad de la figura motivacional en el plano constitucional, corresponde ahora contrastar esta realidad con la siguiente sección que explora la materia de Inteligencia Artificial a un nivel técnico para determinar el grado de compatibilidad entre motivación y las nuevas herramientas tecnológicas.

6. Inteligencia Artificial

El desarrollo tecnológico ha significado grandes avances en el desempeño humano de ciertas tareas. Gracias a la capacidad de automatización es posible aumentar la productividad y minimizar el margen de error. Este creciente avance trajo consigo un nuevo parte aguas en el campo de la informática y ciencias de la computación, la Inteligencia Artificial. El potencial de esta nueva disciplina ha despertado gran preocupación en distintos sectores, particularmente cuestionando y comparando el desempeño de estos modelos con los procesos cognitivos de un ser humano, surgiendo dudas sobre el método de razonamiento empleado por la máquina para asimilar el proceso intelectual humano.

La presente sección del trabajo se centrará en el estudio y definición de la Inteligencia Artificial, delimitando el campo de estudio, así como indagando en su funcionalidad a nivel

técnico. Para ello, se abordará el funcionamiento interno de los Grandes Modelos de Lenguaje (GML), analizando su arquitectura, su forma de razonamiento denominada “geométrica”, y las implicancias que conlleva este enfoque en términos de comprensión del lenguaje, transparencia y trazabilidad de decisiones. Asimismo, se expondrán los desafíos que plantea la denominada “caja negra” algorítmica y las dificultades inherentes para auditar los procesos de decisión de estos modelos.

Finalmente, se introducirá el concepto de Inteligencia Artificial General (AGI) como un horizonte teórico que aspira a superar las limitaciones estructurales actuales de los GML, abriendo un debate profundo sobre las capacidades cognitivas de las máquinas y su eventual rol en contextos de alta responsabilidad, como el jurídico.

6.1. ¿Qué es la Inteligencia artificial?

El término *Inteligencia Artificial* (en adelante IA) es utilizado para referirse a la capacidad adoptada por los sistemas informáticos de imitar las capacidades cognitivas humana, desde un punto de vista legislativo la Unión Europea ha optado por definirla como un sistema diseñado para desempeñar con distintos autonomía multi gradual y adaptabilidad para el cumplimiento de tareas específicas, generando resultados o predicciones⁴². La definición adoptada por la Unión Europea pretende encasillar de manera pormenorizada un concepto tan complejo como lo es la Inteligencia Artificial.

Dentro de la presente exposición es indispensable reconocer e identificar la funcionalidad y naturaleza de los Grandes Modelos de Lenguaje (En adelante GML), siendo estos una progresión en los avances tecnológicos de sistemas de Inteligencia artificial orientados al procesamiento de lenguaje⁴³. Estos modelos guardan características importantes en su arquitectura que merecen recibir un meticuloso estudio, de modo que a través del entendimiento de su proceso y funcionalidad sea posible discernir la naturaleza del lenguaje generado de manera artificial y contrastarlo con el proceso lógico de un juez humano.

⁴² Artículo 3, Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) (DOUE núm.1689, de 12 de julio de 2024).

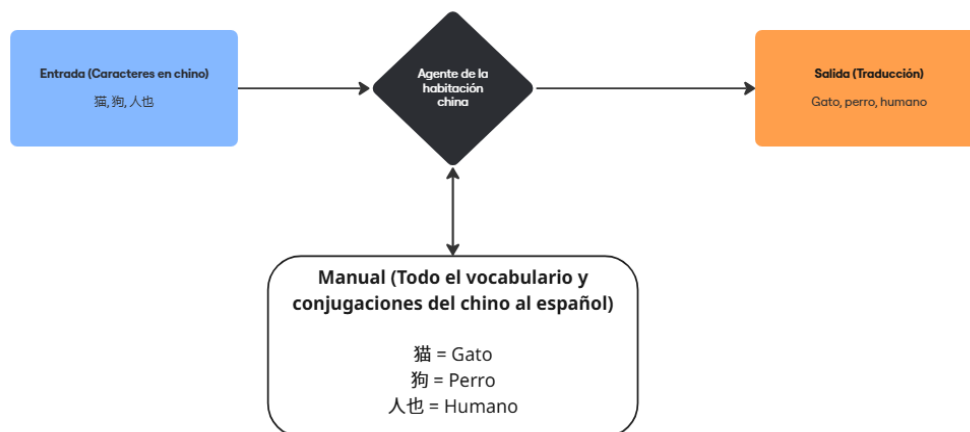
⁴³ Andrea Filippo, Davide Audrito, Luigi Di Caro y Cristina Poncibo, “The architecture of language: Understanding the mechanics behind LLMs”, *Cambridge Forum on AI: Law and Governance* 1, (2025), 2

6.2. Razonamiento de los Grandes Modelos de Lenguaje

El proceso lógico que sigue un GML es fundamentalmente distinto al que sigue un ser humano. A pesar de que su estructura básica busca simular el funcionamiento natural del pensamiento humano, la metodología utilizada para el procesamiento de información es vastamente distinto, pues el entendimiento del lenguaje que posee un ser humano dista sustancialmente del procesamiento que efectúa la máquina.

A modo de ilustrar esta realidad resulta relevante describir el experimento mental desarrollado por Searle en 1980, siendo este el experimento mental de la habitación china. La habitación china describe una situación ficticia en donde una persona que se encuentra dentro de una habitación tiene la tarea de traducir caracteres en chino, el intérprete de este escenario no conoce el idioma, pero cuenta con un manual en su lengua materna que le explica las traducciones específicas de cada símbolo en chino. De este modo es capaz de traducir con precisión el chino para los observadores que se encuentren fuera de la habitación, a pesar de que únicamente realiza acciones mecánicas guiadas por su manual, sin reflejar una verdadera comprensión del chino.⁴⁴ El siguiente gráfico ejemplifica este supuesto:

Gráfico No. 1 Representación gráfico del experimento de la Caja China



Fuente: Elaboración propia en base a fuente bibliográfica.⁴⁵

⁴⁴ Searle, J. R, "Minds, brains, and programs", *Behavioral and Brain Sciences* 3 (1980) 417-418.

⁴⁵ Id.

Searle planteó este escenario para cuestionar los sistemas de la época y la noción de una posible Inteligencia Artificial que pueda realmente acercarse a la inteligencia humana. Si bien es evidente que existe una gran diferencia de desempeño entre los sistemas de IA de finales del siglo XX con los modelos contemporáneos, el argumento se mantiene vigente dado que independientemente del progreso y mejora de los resultados generados, la técnica empleada arrastra el problema trascendental descrito por la habitación china, y es que la mera manipulación de símbolos responde únicamente a las reglas formales o sintaxis del lenguaje, pero ello no implica el acceso al campo semántico de las palabras, quedando limitados al formalismo de la estructura del lenguaje pero sin poder acceder a su significado.⁴⁶

En la actualidad los GML son diseñados en base de la arquitectura del *Transformer*, introducida en 2017 y que significa un paso agigantado en el desarrollo tecnológico, puesto que deshecha la tendencia de utilizar redes neuronales y en su lugar propone un sistema de atención que le facilita el entendimiento de información secuencial de manera íntegra y simultánea.⁴⁷ Es importante mencionar que esta arquitectura no es un entendimiento capaz de generar abstracciones como el de un ser humano⁴⁸, en su lugar, los GML tienen un razonamiento denominado como “geométrico”.⁴⁹

El razonamiento geométrico implica la transformación del lenguaje a un idioma que la máquina pueda comprender, por lo que plasma los elementos de la sintaxis del lenguaje y las procesa a través de representaciones gráficas en un plano matemático.⁵⁰ Esta forma de “comprensión” que tienen los GML es desde una perspectiva del cálculo frío, donde se extrapolan las cualidades del lenguaje natural y lo despojan de su finalidad comunicativa. Dado que este tipo de razonamiento geométrico no permite la abstracción conceptual de los términos del lenguaje, se pone en cuestionamiento la verdadera comprensión realizada por estos modelos. En otras palabras, las máquinas no comprenden conceptos como lo hacen los seres humanos.⁵¹

⁴⁶ Searle, J. R, “Minds, brains, and programs”, *Behavioral and Brain Sciences* 3, 422

⁴⁷ Andrea Filippo, Davide Audrito, Luigi Di Caro y Cristina Poncibo, “The architecture of language: Understanding the mechanics behind LLMs”, *Cambridge Forum on AI: Law and Governance* 1,(2025), 2.

⁴⁸ Alberto Vélez León, “El entendimiento humano”, *Iustitia Socialis* 4 (2019), 41-42.

⁴⁹ Venkat Venkatasubramanian, “Do large language models ‘understand’ their knowledge?”, *AICHe Journal* 71 (2024), 4.

⁵⁰ Id, 5-7.

⁵¹ Venkat Venkatasubramanian, “Do large language models ‘understand’ their knowledge?”, 7.

6.3 Transparencia y el problema de la caja negra

Una vez realizado el análisis del proceso de razonamiento efectuado por un GML y determinado que, por su arquitectura actual son incapaces de comprender plenamente el lenguaje, resulta particularmente relevante detallar un problema fundamental que guardan estos sistemas y que encuentran cobran especial relevancia para el ámbito jurídico, el problema de caja negra que acarrear los GML.⁵² El término de “caja negra” hace analogía al registrador de vuelo presente en aviones, donde estos contienen toda la información pertinente al vehículo, y no es hasta que visualizamos su contenido que podemos realmente conocer el estado y las condiciones que presentaba el avión en situaciones de interés, como en un accidente.⁵³ Así se reconoce que un sistema es una caja negra cuando sus mecanismos internos son desconocidos para el usuario, volviendo inaccesible para el usuario determinar la manera en que la máquina llegó a una respuesta a partir de la interacción inicial.⁵⁴

Tal como se expuso, el procedimiento lógico y de procesamiento de lenguaje es un cálculo mecánico que busca generar predicciones que se adecúen a las estructuras formales de las palabras. Sin embargo, determinar con precisión la lógica detrás de este proceso presenta una serie de dificultades de transparencia, Vera y Wortman Identifican 3 maneras en que el problema de transparencia se manifiesta: 1) Opacidad de los modelos; 2) Limitado acceso a datos de entrenamiento; 3) Modelos propietarios de código cerrado.⁵⁵ Cada una de estas limitantes denotan la dificultad de contar con sistemas de IA que sean plenamente transparentes con el usuario. A continuación, se realizará una breve reseña de cada problema destacado.

La opacidad hace referencia a la dificultad que tienen los usuarios para entender cómo un algoritmo llegó a su decisión en función de la instrucción dada por el usuario.⁵⁶ La

⁵² Vera Liao, Jennifer Wortman Vaughan, “AI Transparency in the Age of LLMs: A Human Centered Research Roadmap”, *Harvard Data Science Review*, 3.

⁵³ Fleming Rutherford, “Dentro de la caja negra: reflexiones sobre el futuro del análisis de la conducta en el siglo XXI”, *Revista Mexicana de Psicología* 25 (2008), 59-70.

⁵⁴ Marcin Rabiza, “A Mechanistic Explanatory Strategy for XAI”, *Philosophy of Artificial Intelligence: The State of the Art*, (2023), 1-2.

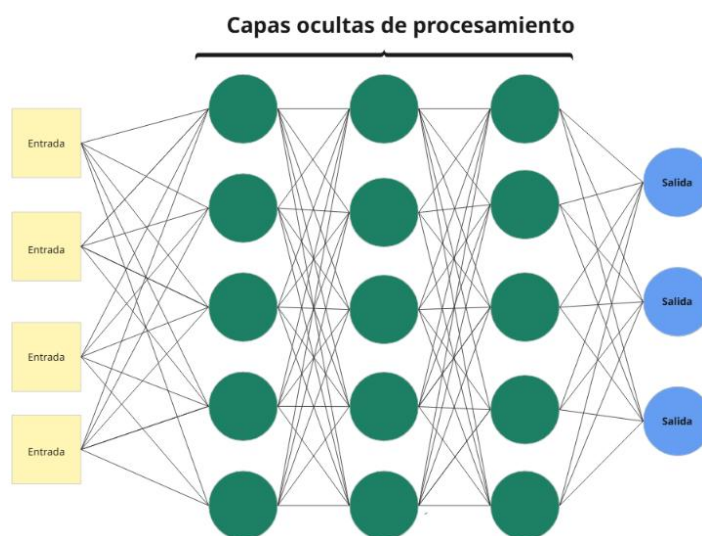
⁵⁵ Id, 6-9

⁵⁶ Ignacio Cofone, “AI and Judicial Decision-Making”, 10

opacidad constituye un problema fundamental en la determinación del camino lógico adoptado por un algoritmo, por lo que realizar un control de arbitrariedad resulta sumamente complejo, pues la propia arquitectura de los Grandes Modelos de Lenguaje sigue siendo compleja de comprender para expertos, dado que el comportamiento del algoritmo y las técnicas de interpretación de este no son precisas, implicando que incluso los expertos tienen dificultades para auditar a ciencia cierta el proceso lógico de un GML.⁵⁷

Dentro del funcionamiento base de una red neuronal, esta comprende capas la información de entrada se conecta a neuronas ocultas, estas capas son escondidas del usuario y realizan el cálculo en base a la información de entrada proporcionada por el usuario.⁵⁸ Por lo que el usuario solo puede apreciar la capa de salida, es decir, la respuesta otorgada por la máquina, sin acceder en ningún momento al razonamiento que se efectuó para llegar a esa respuesta. La siguiente gráfica identifica la estructura simple de una red neuronal, además permite visualizar la opacidad presente en sus neuronas ocultas.⁵⁹:

Gráfico No. 2 Modelo de Red Neuronal y sus capas ocultas



Fuente: Elaboración propia en base a fuente bibliográfica

⁵⁷ Tolga Bolukbasi, et al. "An Interpretability Illusion for BERT", 8.

⁵⁸ Chul-Heung Kim, Suhwan Lim, Sung Yun Woo , Won-Mook Kang ,Young-Tak Seo , Sung-Tae Lee , Soochang Lee , Dongseok Kwon ,Seongbin Oh , Yoohyun Noh , Hyeongsu Kim , Jangsaeng Kim ,Jong-Ho Bae and Jong-Ho Lee, "Emerging memory technologies for neuromorphic computing", *Nanotechnology* 30, (2019), 15.

⁵⁹ La red neuronal en específico a la que se hace referencia es la red neural profunda, perteneciente al modelo de *Deep Learning* que es uno de los enfoques más relevantes en el desarrollo de la Inteligencia Artificial en la era moderna.

La segunda barrera se refiere a los datos de entrenamiento con el que un GML cuenta en su base para esquematizar sus parámetros internos, los modelos de IA son provistos de grandes cantidades de información que conforman la base de su conocimiento.⁶⁰ No obstante, esta información usualmente queda resguardada bajo secreto empresarial de las entidades desarrolladoras de un GML, impidiendo al público general conocer los datos sobre los que está razonando el algoritmo,⁶¹ y dificultando la determinación de posibles sesgos inherentes al tipo de datos empleados en el entrenamiento del modelo.

Finalmente, el problema de los modelos propietarios de código cerrado abarca un conjunto de complicaciones relativas a los aspectos técnicos del modelo, particularmente señalando el código fuente del algoritmo, sus parámetros, sus pesos internos, entre otros.⁶² Si las características señaladas son muy relevantes para el correcto entendimiento de un modelo en particular, considero que el aspecto más importante para la determinación transparente del razonamiento de un GML se encuentra en la publicidad de sus pesos internos, más allá que el propio código.

Los pesos son magnitudes matemáticas utilizadas en el cómputo de probabilidades, acceder a estos pesos implica acceder al cálculo que llevó a un modelo tomar una decisión respecto de la orden recibida.⁶³ Puesto en palabras simples, los pesos determinan qué palabra tiene más probabilidades de aparecer en un contexto determinado. Sin embargo, estos ajustes son tan complejos que ni siquiera los desarrolladores pueden predecir con exactitud cómo un modelo llega a una respuesta.

Por otro lado, la opción de publicitar el código fuente y hacerlo abierto al público (referido en el rubro como *open sourcing* y *open weights*) es comparativamente menos eficiente si se busca garantizar plena transparencia, puesto que el código es, de manera simplificada, la maquinaria que permite al sistema realizar una tarea específica, pero no es el medio idóneo para visualizar el proceso.⁶⁴

⁶⁰ Andrea Filippo et al. “The architecture of language: Understanding the mechanics behind LLMs”, 3.

⁶¹ Ignacio Cofone, “AI and Judicial Decision-Making”, 11.

⁶² Vera Liao, Jennifer Wortman Vaughan, “AI Transparency in the Age of LLMs: A Human Centered Research Roadmap”, 7.

⁶³ Andrea Filippo et al. “The architecture of language: Understanding the mechanics behind LLMs”, 2.

⁶⁴ Stefan Larson, Fredrik Heintz, “Transparency in Artificial Intelligence”, *Internet Policy Review* 9, (2020), 6-7.

Superar estas barreras es crucial para fomentar un esquema de Inteligencia Artificial que sea plenamente transparente en su toma de decisiones. Sin embargo, queda remanente una problemática inherente a estos sistemas, si bien se identificó los puntos a publicitar por parte de los sistemas, estas características sufren de un problema de interpretabilidad y explicabilidad general. La interpretabilidad se refiere a la capacidad de entender el funcionamiento de un modelo y la explicabilidad es el proceso que el modelo mismo proporciona al usuario para explicar su funcionamiento.⁶⁵ Esto es así dado que las distribuciones y cálculos vectoriales efectuados por el algoritmo son por sí mismas difíciles de determinar inclusive para expertos en el campo, creando una aparente pared infranqueable en la búsqueda de establecer un marco de transparencia.

Así es como varios especialistas plantearon el ideal de una Inteligencia Artificial explicable (en adelante XAI), este estándar busca someter a la IA en general a la adopción de modelos plenamente transparentes cuyo proceso de toma de decisiones sea transparente para el usuario en su interpretabilidad y cuente con transparencia algorítmica que facilite la comprensibilidad matemática para los expertos.⁶⁶ Dentro del campo de la XAI se han desarrollado técnicas encaminadas a suplementar la opacidad, tales como los Modelos de explicación locales (LIME) y Shapley values.⁶⁷

La transparencia es un aspecto de suma importancia dentro de los GML, pues un marco que permita a los usuarios conocer el camino lógico empleado es de vital importancia cuando se traduce su uso al ámbito jurídico. Si bien se han planteado técnicas en busca de elevar el estándar de transparencia adoptada por los sistemas, estas aún no terminan de solventar las carencias de trazabilidad decisional que son características de los GML.

6.4. Inteligencia Artificial de carácter General

Habiendo indagado en las propuestas de transparencia implementada a los sistemas de Inteligencia artificial, es igualmente relevante indagar en el concepto denominado como

⁶⁵ Alejandro Barredo, Natalia Díaz, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil, Daniel Molina, Richard Benjamins, Raja Chatila, Francisco Herrera, “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI”, *Information Fusion*, 58, (2019), 6.

⁶⁶ Alejandro Barredo, et al. “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI”, 10-11.

⁶⁷ Id, 18.

“Inteligencia Artificial General” (en adelante AGI por sus siglas en inglés) y la manera en que este modelo teórico podría ser una solución a futuro que logre solventar los problemas estructurales que acarrearán actualmente los GML.

El postulado de la Inteligencia Artificial General apuesta por un nivel de Inteligencia Artificial que alcanza la generalidad cognitiva que le permita desempeñarse de manera óptima en múltiples tareas, y que además represente un nivel de adaptabilidad que esté fuera de la expectativa de sus creadores.⁶⁸ La precisión y calidad de las actuaciones cognitivas desarrolladas por un sistema de AGI sería comparable o superior al de un ser humano, dependiendo del nivel de sofisticación del sistema.⁶⁹

Para determinar la viabilidad de la AGI como una respuesta a futuro para la emisión de decisiones que no incurran en los errores de actuales modelos de IA⁷⁰, es imprescindible comprender los principios claves que el planteamiento ideal de una AGI representa, siendo este propuesto por Bostrom, donde define a una AGI como un sistema diseñado con la capacidad de aprender autónomamente, razonar bajo incertidumbre, formar y combinar conceptos flexiblemente, y realizar cualquier tarea cognitiva que un humano pueda hacer,⁷¹ todo ello con potencial de mejora progresiva.⁷²

Si bien todos los principios enunciados guardan su particular relevancia para la discusión global de una AGI, para la presente exposición me enfocaré en explorar los primeros 3 puntos que considero de mayor relevancia para el tema central de este trabajo. Sobre el enfoque en las capacidades, los autores defienden la idea que una AGI no se ve definida por su metodología de funcionamiento, sino que las capacidades que alcanza son

⁶⁸ Ben Goertzel, “Artificial General Intelligence: Concept, State of the Art, and Future Prospects”, *Journal of Artificial General Intelligence*, 5, (2014), 2

⁶⁹ Meredith Ringel Morris, Jascha Sohl-Dickstein, Noah Fiedel, Tris Warkentin, Allan Dafoe, Aleksandra Faust, Clement Farabet3 Shane Legg, “Position: Levels of AGI for Operationalizing Progress on the Path to AGI”, arXiv, (2023), 5

⁷⁰ Particularmente hablando sobre sesgos presentes en las decisiones que se encuentren influenciados por sus datos de entrenamiento; Así como la naturaleza de su razonamiento que se basa puramente en el cómputo matemático y la distribución de probabilidades.

⁷¹ Este postulado ha sido desarrollado, de modo que actualmente existe la corriente que considera que la AGI se debe enfocar netamente en los procesos cognitivos, distinguiendo seis puntos clave para este propósito: 1) Enfoque en capacidad y no procesos; 2) Enfoque en Generalidad y rendimiento; 3) Enfoque en habilidades Cognitivas y metacognitivas; 4) Enfoque de potencial; 5) Enfoque en sostenimiento ecológico; 6) Enfoque en la AGI como un camino.

⁷² Nick Bostrom, *Super Intelligence*, (Londres: Oxford University Press, 2014), 51-53.

verdaderamente lo que lo dota de valor.⁷³ De este modo si lo trasladamos al campo de los GML implica que dejamos de definir al sistema por medio de una arquitectura específica que delimita la manera en que funciona, el postulado de una Inteligencia Artificial General trasciende la necesidad de encasillarla por su diseño y se enfoca netamente en la calidad de resultados que forja.

El segundo punto importante para tratar versa sobre el enfoque cognoscitivo general, este hace hincapié en que una AGI debe ser capaz de realizar una generalidad de tareas que no impliquen reprogramación ni adaptación externa⁷⁴, esencialmente permitiendo alejarse del problema de sesgos presentes en los algoritmos actuales y, abriendo la posibilidad que los sistemas de manera autónoma puedan regular sus propios parámetros para identificar y mitigar la presencia de razonamientos sesgados. Esta característica en particular en conjunto con respuestas sofisticadas que repliquen el criterio de un experto reviste a este ideal de Inteligencia General de plenitud cognitiva que actualmente no poseen los sistemas de IA convencionales.

Finalmente, el enfoque de una AGI deberá centrarse en el desarrollo de habilidades cognitivas, es decir, que su inteligencia sea medida usando parámetros de inteligencia cognitiva y descartando las habilidades físicas⁷⁵ que puedan llegar a tener.⁷⁶ La ventaja de esta aproximación es que permite el desarrollo pleno de IA basándose únicamente en la perspectiva del procesamiento y razonamiento de la información, pues incursionar en el campo de la robótica es intrascendente desde el punto de vista de la evolución neta de la inteligencia.

Por consiguiente, el objetivo de alcanzar una eventual AGI abre las puertas a que la discusión sobre decisiones plenamente motivadas pueda desarrollarse aún más, considerando que la hipotética AGI se deslinda de la rigidez parametrizada de la IA actual y se convierte en una fuente de conocimiento que desempeña habilidades cognitivas que realmente se asimilan a las decisiones humanas, sin contar con limitaciones de seguir patrones y esquemas internalizados y pudiendo adaptarse a nuevas circunstancias. Cabe destacar que el marco de

⁷³ Nick Bostrom, *Super Intelligence*, 53.

⁷⁴ Meredith Morris et al., “Position: Levels of AGI for Operationalizing Progress on the Path to AGI”, 4.

⁷⁵ Refiriéndose al campo de la robótica como medio para dotar a los sistemas de IA de una manera de interactuar con el mundo físico.

⁷⁶ Meredith Morris et al., “Position: Levels of AGI for Operationalizing Progress on the Path to AGI”, 6.

la AGI constituye un ideal teórico que, al momento de la redacción de este trabajo, aún no ha sido alcanzado pero que no deja de ser un escenario plausible a futuro.

7. Capacidad motivacional de Grandes modelos de Lenguaje

Una vez estudiadas las teorías motivacionales como el funcionamiento de los GML desde una perspectiva técnica, la presente sección tiene el objetivo de analizar la capacidad que posee la máquina para emitir decisiones que gocen de suficiencia motivacional a la luz del Derecho, verificando que cumpla tanto con los aspectos lógicos, epistemológicos y axiológicos.

7.1. ¿En qué se diferencia el entendimiento humano del artificial?

La discusión sobre la forma en que los seres humanos experimentan el mundo y adquieren conocimiento ha sido objeto de debate y desarrollo durante siglos. Durante el desarrollo de la historia, se ha teorizado de diversas maneras sobre el origen y la naturaleza del intelecto humano, algunas posturas lo han vinculado con conceptos esotéricos, asociándolo con la existencia de un alma y valores teológicos.⁷⁷

Dentro del estado del arte de las acepciones académicas sobre el entendimiento humano, surgen postulados de relevancia como el propuesto por Popper que sostiene que el entender implica la formación de conjeturas, analizarlas y aprender de los errores;⁷⁸ por su parte otro de los postulados de mayor relevancia es aquel formulado por Sapolsky que posee una visión biológicamente determinista, considerando que la capacidad crítica y analítica se somete al contexto biológico del ser humano.⁷⁹

Trasladando esta idea hacia los GML encontramos un problema madre respecto a una visión empírica del entendimiento, siendo esta el hecho que los sistemas actuales que hacen uso de la arquitectura del *transformer* no comprenden el lenguaje de la misma manera que un ser humano, pues no son capaces de la abstracción conceptual y su funcionamiento se basa en predicciones. Remitiéndome al experimento de la habitación china, así como el agente

⁷⁷ Manuel Correia, “La naturaleza del intelecto humano en el De principiis de Orígenes de Alejandría”, *Teología y vida* 55, (2017), 49.

⁷⁸ Ver Karl Popper, *El Cuerpo y la Mente* (Barcelona: Ediciones Paidós, 1997), 31-84

⁷⁹ Ver Robert Sapolsky, *Decidid: Una Ciencia de la vida sin libre albedrío* (Capitán Swing Libros, 2024), 228.

interpretador dentro de la habitación no comprende realmente el chino, el GML tampoco entiende las palabras, ya que únicamente manipula los elementos morfológicos y sintácticos del lenguaje, sin llegar a procesar su valor semántico.

Sin embargo, Agüeras desafía esta idea al sostener que las diferencias existentes entre el entender humano y el entender de las máquinas son irrelevantes para enmarcar las capacidades intelectuales de la IA, formulando que el reduccionismo estadístico del proceso lógico efectuado por los GML no debe verse como limitante solo por distar de la realidad humana, sino que debe verse paralelamente como un nuevo paradigma del entendimiento característico de nuevas y distintas mentes no biológicas. Desde esta perspectiva, el funcionamiento estadístico constituye realmente una construcción de significado, alegando que la denominada “comprensión real” no se diferencia del proceso predictivo, sino que ambas representan perspectivas alternativas sobre el mismo fenómeno.⁸⁰

Esta segunda tesis, a pesar de abrir un debate interesante, no termina de resolver la limitante más importante que tienen los GML en comparación al entendimiento, siendo este el carácter sintáctico y no semántico de sus procesos. Por más que los resultados generados por estos modelos puedan simular comprensión, su actividad se reduce a manipulaciones de símbolos guiadas por relaciones de la matemática, sin acceso alguno al significado de los conceptos que operan, al igual que en la habitación china.⁸¹

7.2 ¿Puede un Gran Modelo de Lenguaje motivar?

Así como se indagó en la primera sección; la motivación no se limita a una simple exposición de ideas o razones lógicas que persigan un silogismo, sino que atiende a un deber superior ligado a la legitimación del poder judicial y el escrutinio de la sociedad. Bajo esta óptica, queda claro que se requiere más que la capacidad de hilar justificaciones que mantengan un curso lógico. Siguiendo esta premisa, surge la interrogante respecto a si los GML poseen realmente la capacidad de emitir decisiones que gocen de una verdadera motivación.

⁸⁰ Blaise Agüeras, “Do Large Language Models Understand Us?”, *Daedalus* 151, (2022), 190-193.

⁸¹ Esto es considerando la tecnología actual, y dejando abierta la posibilidad que una eventual AGI posea estas habilidades y avive el debate.

Para abordar esta pregunta, sugiero enfocar la presente sección en dos puntos principales de análisis: 1) Cumplimiento del fin motivacional; 2) Cumplimiento de la lógica jurídica. Tratando el primer criterio a analizar, reconocemos que el fin de la institución motivacional contemporánea es la de ofrecer a la ciudadanía un mecanismo de control respecto al uso del poder público, permitiendo tanto a las partes procesales como a la población externa conocer las razones que conllevaron a una resolución, a manera de evitar un sistema de justicia arbitrario.⁸² Frente a esto, un GML presenta limitantes muy marcadas que dificultan la plena garantía de transparencia en sus respuestas, partiendo del hecho que los métodos de publicitar el modelo empleado, tales como los datos de entrenamiento y el código fuente, son insuficientes para determinar el proceso de razonamiento efectuado.

A pesar de existir formas de mitigar esta problemática por medio de técnicas de explicabilidad e interpretabilidades propuestas por la corriente de la XAI,⁸³ el diseño de los sistemas y su manera de razonar complican enormemente la transparencia de su proceso justificativo, imposibilitando a los usuarios conocer a cabalidad el verdadero camino racional adoptado por un GML para alcanzar una respuesta. Pues, como se definió con anterioridad, la motivación no examina la calidad o correctitud en sí, sino el proceso justificativo que conllevan estas decisiones.⁸⁴

El segundo punto se centra en el cumplimiento de la lógica jurídica, pues la faceta lógica de la institución motivacional tiene una relevancia notoria en la formulación de esquemas racionales que persigan un camino coherente. Así, uno de los atributos esenciales del razonamiento judicial es la capacidad deductiva del juzgador, MacCormick destaca la capacidad deductiva del juez en su capacidad de armar un esquema lógico mediante la aplicación de reglas inferenciales válidas.⁸⁵ La mencionada capacidad deductiva siendo una cualidad propia de la aplicación lógica humana que, si bien puede ser hasta cierto punto replicable por un GML, es una simulación resultante de un proceso abductivo carente de valoraciones axiológicas y epistémicas, marcando una importante diferencia entre el intelecto

⁸² Michele Taruffo, *Proceso y Decisión Lecciones mexicanas de Derecho Procesal*, (Madrid: Marcial Pons, 2012), 108

⁸³ Entendiéndose por XAI la rama comentada que busca dotar a los sistemas de IA de una mejor trazabilidad y explicabilidad, de modo que pueda establecer un marco de transparencia íntegro.

⁸⁴ Manuel Atienza, *Curso de argumentación jurídica*, 133-134.

⁸⁵ Ver Neil MacCormick, *Razonamiento Jurídico y Teoría del Derecho*, (Lima: Palestra editores, 2018), 49-56

humano y la capacidad actual de los sistemas⁸⁶ que no compagina con la metodología humana de aplicar axiomas lógicos.

Ante esta postura puede surgir la interrogante respecto a la justiciabilidad de una razón emitida por un GML, pues desde la perspectiva del usuario, las respuestas obtenidas podrían ajustarse bajo esquemas fácticos y normativos que persigan un hilo lógico independientemente de que la metodología sea distinta a la humana. Si bien es una perspectiva válida, es importante considerar que los resultados generados por un GML pretenden ser una simulación de las capacidades cognitivas humanas, y esta naturaleza simulada del razonamiento termina por comprometer la integridad axiológica de la motivación, al no comprender ni efectuar un verdadero juicio valorativo.

Estas corrientes dejan en evidencia que las diferencias entre el entendimiento humano y el entendimiento artificial marcan dificultades en el acto resolutorio desde la práctica. Pues, los GML carecen de cualidades humanas, siendo incapaces de comprender plenamente el lenguaje natural desde el enfoque semántico, efectuando exclusivamente la manipulación estadística de las formas para generar secuencias de texto comprensibles para el usuario, pero que no implican una internalización ni abstracción conceptual del lenguaje.⁸⁷ Esta deficiencia en el juicio contextual pone en duda la capacidad de los sistemas de generar una motivación suficiente, pues se incurre en un problema tanto práctico como hermenéutico, pues la aplicación generada por GML dista sustancialmente de las características que revisten una resolución emitida por un humano.

Adicionalmente, es crucial reconocer que la aparente racionalidad exhibida por un GML opera exclusivamente en un plano descriptivo. Es decir, su razonamiento proviene de la identificación de patrones en el lenguaje, generando así respuestas que caen en la mera descripción. Esta descriptividad conceptual es insuficiente para fundamentar su validez normativa, pues no identifica ni atribuye valor a las reglas generales. Pretender que la manera en que un GML es determina la manera en que debe ser una justificación jurídica válida, o que su resultado constituye per se una razón legítima, implicaría incurrir en la falacia

⁸⁶ Vera Liao, et al., “AI Transparency in the Age of LLMs: A Human Centered Research Roadmap”, 5.

⁸⁷ Ver Tolga Bolukbasi, et al. “An Interpretability Illusion for BERT”, 7-8

naturalista donde se intenta acceder a la esfera valorativa desde una perspectiva descriptiva que es incapaz de conocerla.⁸⁸

Si bien los GML pueden generar texto que simule una motivación judicial, carecen de los elementos esenciales que definen este proceso en el derecho: Integridad axiológica, la habilidad deductiva, la transparencia y la integración de valores normativos.⁸⁹ Como se ha demostrado en este análisis, la motivación judicial va mucho más allá de la aplicación de patrones lógicos o lingüísticos, ya que implica una actividad interpretativa y justificativa que solo un agente humano puede realizar de manera plena.

7.3 Incapacidad axiológica de los Grandes Modelos de Lenguaje

Siguiendo la línea argumentativa desarrollada en las secciones precedentes, es imperativo reconocer que la motivación judicial trasciende la mera aplicación silogística de normas o la construcción de cadenas lógicas formalmente coherentes.⁹⁰ Su propósito fundamental radica en el resguardo de los principios y valores que sustentan un Estado constitucional de derecho, actuando como un mecanismo esencial para prevenir la arbitrariedad en el ejercicio del poder jurisdiccional.⁹¹ La justificación de una decisión judicial no es, por consiguiente, un ejercicio puramente mecánico o lógico-formal, sino un acto también impregnado de consideraciones axiológicas.⁹²⁹³

Bajo este principio, los GML presentan una limitación intrínseca y fundamental. Su funcionamiento, como se ha evidenciado, se encuentran estrechamente ligados al análisis estadístico de la vasta cantidad de información con la que fueron entrenados y corre el riesgo inherente de reproducir respuestas sesgadas.⁹⁴ Esta dependencia de patrones preexistentes

⁸⁸ Orlando de la Vega, “Falacia Naturalista y Ley De Hume: Su Significado En Derecho Penal”, *Revista chilena de Derecho* 47, 2020, 58-59.

⁸⁹ Gastón Valenzuela, “Enfoque actual de la motivación de las sentencias. Su análisis como componente del debido proceso”, *Revista de Derecho Universidad de Montevideo* 21, (2020), 82

⁹⁰ Piero Calamandrei, *La motivación de la sentencia Civil*, 125.

⁹¹ Michele Taruffo, *Proceso y Decisión Lecciones mexicanas de Derecho Procesal*, 94.

⁹² Manuel Atienza, *Curso de argumentación jurídica*, (Madrid, Trotta, 2013), 133-134.

⁹³ Dentro de las corrientes del Derecho es debatida la influencia de los valores morales en la validez legal. Desde una perspectiva iusnaturalista el derecho y la moral mantienen una relación intrínseca, mientras que para el positivismo excluyente se niega esta relación para determinar la validez del Derecho; Dentro del positivismo jurídico también existe la corriente incluyente que reconoce la influencia moral, pero la observa desde una posición contingente que no debe influir en mayor medida en la aplicación del Derecho.

⁹⁴ Andrea Filippio, et al., “The architecture of language: Understanding the mechanics behind LLMs”, 14-15.

puede conducir a la formulación de razonamientos que, aunque puedan parecer formalmente coherentes, operan bajo un esquema lineal y predictivo, incapaz de realizar la precisa ponderación valorativa que exige la administración de justicia en muchos casos.

Las respuestas sesgadas generadas por estos sistemas responden tanto a la calidad como a la variedad de datos con los que fueron entrenados, Cofone identifica tres sesgos principales⁹⁵: 1) Sesgo dentro del proceso algorítmico⁹⁶; 2) Sesgo no representativo⁹⁷; 3) Sesgo social arraigado.⁹⁸ Los sesgos replicados por los sistemas son particularmente problemáticos cuando se traducen a la práctica jurisdiccional, pues definen en gran medida el camino por el que transitarán las decisiones en una línea general.⁹⁹ Mientras que no es fuera de lo común que un juzgador se aleje de una marcada tendencia jurisprudencial, los GML que dependen de detectar y seguir patrones de su información almacenada, seguirán la corriente marcada por los datos que comprendan sus parámetros internos, potencialmente llevando a perpetuar acciones discriminatorias¹⁰⁰ o bien, evitar la apertura de nuevas posturas jurisprudenciales que enriquezcan el arte de juzgar.

Más allá de la problemática de los sesgos derivados de sus datos de entrenamiento, subyace una imposibilidad quizás más profunda, la incapacidad estructural de los GML para replicar la dimensión valorativa que caracteriza a la justificación externa de una sentencia. Como la teoría de la argumentación jurídica ha puesto de relieve, la decisión judicial requiere que el juzgador no solo demuestre la coherencia lógica de su razonamiento, sino que también fundamente su fallo en su adecuación a los principios, derechos y valores fundamentales que informan el ordenamiento jurídico y que son reconocidos por la sociedad.¹⁰¹ Un GML, en virtud de su naturaleza computacional, no puede simular ni incorporar las cualidades intrínsecamente humanas que nutren este proceso tales como la empatía, la prudencia o la conciencia moral.

⁹⁵ Ver Ignacion Cofone, “AI and Judicial Decision-Making”, 9-10.

⁹⁶ Se refiere a aquellos sesgos arraigados de la interacción humana, es decir, sesgos otorgados por el usuario durante la interacción con el sistema.

⁹⁷ Son aquellos sesgos derivados de la calidad de información de entrenamiento, donde estos datos no representan adecuadamente a la población general.

⁹⁸ Estos sesgos provienen de datos que si son representativos de la población, pero que reflejan del mismo modo la visión discriminatoria presente.

⁹⁹ Ignacion Cofone, “AI and Judicial Decision-Making”, 8.

¹⁰⁰ Id, 10.

¹⁰¹ Juan Igartura Salaverría, *El razonamiento en las decisiones judiciales*, 30

Fuller dentro del texto clásico: “El caso de los exploradores de las cavernas” plantea un escenario hipotético en donde distintos jueces se encuentran en conflicto respecto a un caso que presenta circunstancias extremas y de difícil determinación.¹⁰² Schauer dentro de su análisis explora la manera en que incluso ante una norma clara y una estructura legal bien definida, la toma de decisiones judiciales revela su complejidad al enfrentar dilemas donde la aplicación literal del derecho se envuelve con consideraciones morales, políticas o de sentido común. Estas condiciones son ajenas a la literalidad de un hecho fáctico o de la valoración rígida de la norma, marcando la relación entre los valores sociales con la aplicación del derecho para el desarrollo judicial.¹⁰³ Los sistemas actuales, al no poseer una consciencia meta cognitiva¹⁰⁴ que le permita reconocer y adquirir capacidad de valoración moral, les impide acceder al campo de ajuste axiológico que resulta indispensable para que una decisión pueda considerarse plenamente motivada y justificada desde una perspectiva sustancial.¹⁰⁵

Por consiguiente, la dimensión axiológica de la motivación judicial representa una barrera por ahora infranqueable, para los GML. Su incapacidad para internalizar y aplicar valores, realizar ponderaciones morales complejas y comprender el sustrato ético y social del derecho les impide generar decisiones que satisfagan cabalmente la exigencia constitucional y democrática de una justificación plena, una justificación que vaya más allá de la mera corrección formal o la predicción estadística.

7.4 Comparativa decisiones judiciales reales y artificiales

Habiendo delimitado las dificultades y limitaciones que tienen los GML para la formulación de decisiones motivadas desde un punto de vista teórico, esta última sección se centrará en el análisis llevado a cabo en la Universidad de Osaka, Japón por Watamura et al. Este estudio es un análisis comparativo entre las decisiones emitidas por jueces humanos y las resoluciones judiciales formuladas por GML. Un grupo de personas con limitada

¹⁰² Ver Lon Fuller, *El caso de los exploradores de cavernas*, (Buenos Aires: Abeledo-Perrot, 2002).

¹⁰³ Fredrik Schauer, “Fuller’s Fairness: ‘The case of the speluncean explorers’”, *University of Queensland Law Journal* 35, (2016),

¹⁰⁴ Jack Lindsey, et al., “On the Biology of a Large Language Model”, *Transformer Circuits*, 2025.

¹⁰⁵ Id, 36

experiencia en litigio actuó como jurado, y se encargaron de evaluar ambos tipos de resoluciones sin conocer su procedencia, de modo que podrían elegir la decisión con la que sientan mayor afinidad sin el prejuicio de su procedencia.¹⁰⁶

Para este estudio, se analizaron casos criminales, algunos de los cuales contenían circunstancias atenuantes, mientras que otros eran resoluciones más directas y normativamente claras. Los hallazgos principales del estudio indicaron que, Contrario a la hipótesis de los investigadores (que esperaban mayor deferencia al juez humano en casos con atenuantes) y a hallazgos previos sobre "aversión al algoritmo", los resultados de ambos experimentos no mostraron diferencias significativas en la deferencia de los participantes hacia las recomendaciones del juez humano o de la IA.¹⁰⁷ Es decir, la sentencia final escogida por los participantes no varió significativamente si la recomendación de prisión provenía del juez o de un sistema.¹⁰⁸

El factor que sí tuvo un impacto determinante y significativo en la decisión de los participantes fue la presencia de circunstancias atenuantes, las cuales consistentemente llevaron a sentencias más leves, independientemente de si la recomendación conflictiva provenía del juez o de una máquina.¹⁰⁹ Esto sugiere que los participantes sí ponderaron los hechos específicos del caso, pero su evaluación de la fuente de la recomendación (humana vs. artificial) no alteró su juicio final sobre la pena.¹¹⁰ Los resultados generales del experimento se pueden visualizar en la siguiente gráfica:

¹⁰⁶ Eiichiro Watamura, Yichen Liu, Tomohiro Ioku, "Judges Versus Artificial Intelligence in Juror Decision-making in Criminal Trials: Evidence From Two Pre-registered Experiments", *PLoS ONE* 20 (1), (2025), 2-3

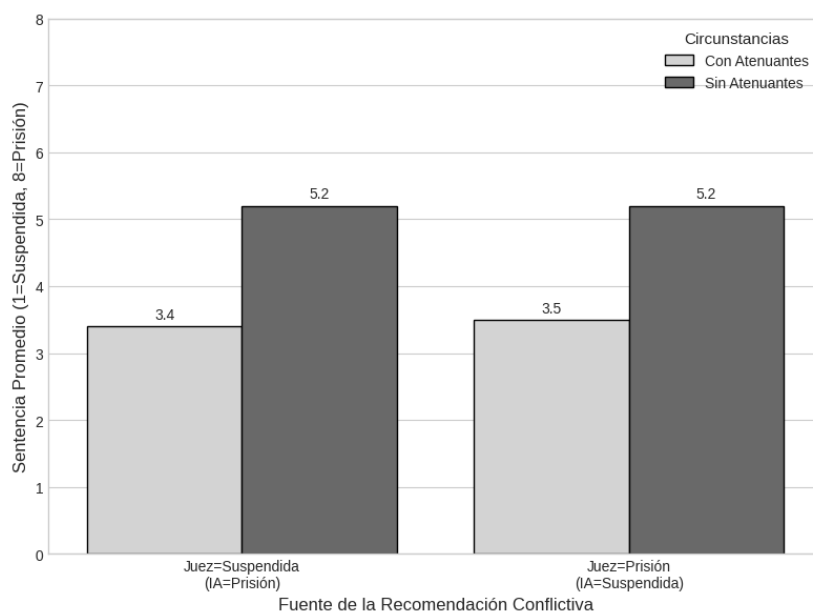
¹⁰⁷ Id, 12

¹⁰⁸ Id, 5

¹⁰⁹ Id, 10

¹¹⁰ Id, 11-13

Gráfico 3. Resultados del experimento



Fuente: Elaboración propia en base a fuente bibliográfica

Si bien el estudio no muestra una diferencia significativa hacia las recomendaciones de jueces y IA en su configuración experimental, esto no implica necesariamente que los GML sean capaces de producir una decisión motivada en el sentido pleno exigido por el Derecho. Como se argumentó previamente, la motivación requiere justificación racional, comprensión semántica, capacidad axiológica y transparencia, elementos que los GML, por diseño, no poseen.¹¹¹ El estudio mide la influencia de una recomendación en una decisión final simplificada, pero no evalúa si los participantes considerasen válida la motivación detrás de la recomendación de la IA si esta fuera expuesta, pues lo que demuestra el estudio es que el algoritmo puede llegar a ser muy eficaz en seguir un patrón judicial marcado que aparente ser humano, pero esta apariencia no responde al fondo valorativo que existe dentro de un fallo judicial.

Si bien este estudio aporta una valiosa perspectiva empírica a un debate predominantemente teórico, los resultados presentados son insuficientes para justificar la transición a un modelo de administración de justicia automatizado, puesto que persisten los problemas intrínsecos de un algoritmo de cara al carácter sustantivo de la motivación.

¹¹¹ Andrea Filippo et al, “The architecture of language: Understanding the mechanics behind LLMs”, 17.

Adicionalmente, es crucial reconocer las limitaciones metodológicas del experimento. En primer lugar, la generalización de los resultados se ve comprometida por la dependencia de una muestra online japonesa, la cual además presentó una alta tasa de exclusión, afectando potencialmente su representatividad.¹¹² En segundo lugar, el contexto cultural y jurídico específico de Japón puede introducir percepciones distintas a las del ámbito occidental o ecuatoriano. Finalmente, la simplificación de la decisión judicial a una escala numérica de 8 puntos, aunque metodológicamente práctica, podría no capturar toda la complejidad del proceso real de sentencia.¹¹³

8. Conclusiones

El presente trabajo ha explorado la institución motivacional, determinando y analizando la viabilidad de adaptarla al campo de la inteligencia artificial generativa mediante Grandes Modelos de Lenguaje. Definiendo a la motivación como un ejercicio justificativo que trasciende el análisis lógico, y que cumple un rol dentro del control del poder público. En este sentido, se deja en evidencia que las decisiones generadas por los GML carecen de las características cognitivas que definen el entendimiento humano, así como presentan problemas de transparencia que imposibilitan cumplir su rol dentro de la esfera constitucional.

Bajo esta premisa se marca el límite que tiene un GML para tomar decisiones que gocen de plenitud motivacional, evidenciando así que sus limitantes inherentes a su diseño imposibilitan la configuración de sus respuestas dentro de la institución motivacional. Tal como se ha mencionado, la motivación no exige únicamente coherencia lógica en el razonamiento del juzgador, sino que también exige que el reflejo de este proceso lógico posea una verdadera comprensión valorativa de los elementos sobre los que pretende resolver. En este sentido, la comprensión semántica y conceptual del lenguaje es algo que las máquinas no pueden replicar más allá de la manipulación formal de la sintaxis.

¹¹² Eiichiro Watamura et al. “Judges Versus Artificial Intelligence in Juror Decision-making in Criminal Trials: Evidence From Two Pre-registered Experiments”, 5

¹¹³ Id, 6-7.

Dentro de los puntos de mayor relevancia que fueron tratados, radica la importancia de la adquisición de conocimiento mediante la experiencia empírica descrita por Locke, esta manera moldea al juzgador y reconoce que el cúmulo de experiencias a lo largo de su vida influye en su capacidad de juzgar y justificar sus decisiones, si bien esta premisa puede ser desafiada en la búsqueda de una administración de justicia objetiva, es una característica del arte de juzgar que guarda valor al momento de decidir.

Finalmente, encarando la importancia de la motivación dentro del estado de derecho los GML presentan un problema de opacidad estructural que imposibilita o dificulta la visión transparente hacia la justificación de sus decisiones. Este problema de transparencia atenta en contra del fin perseguido por la motivación, que es el de servir como una garantía que permita el control del poder judicial y que somete a los administradores de justicia al escrutinio de la sociedad. En este sentido, cuando se analiza la opacidad de los GML desde un punto de vista jurisdiccional ya no es un mero obstáculo técnico, sino que pasa a constituir un problema constitucional al impedir la trazabilidad de las decisiones.

Es por lo previamente expuesto que, el presente trabajo argumenta que los sistemas de IA generativa encargadas del procesamiento de lenguaje natural, a pesar de que aparentan simular los procesos cognitivos que permiten una fundamentación motivacional suficiente, esta aplicación solo se limita a la superficialidad de un esquema lógico que pretende ser plausible. No obstante, esta superficialidad demuestra ser insuficiente en un análisis profundo, pues las barreras asociadas por el diseño de la máquina -tales como la brecha semántica, incapacidad de reconocer y reflejar valores axiológicos, y problemas de transparencia- impiden que esta motivación aparentemente suficiente se encuentre reforzada del verdadero valor sustantivo de la institución motivacional.